

# Vision-based MAV Navigation: Implementation Challenges Towards a Usable System in Real-Life Scenarios

Markus W. Achtelik, Stephan Weiss, Simon Lynen, Margarita Chli, Roland Siegwart  
Autonomous Systems Lab, ETH Zurich, Switzerland

Supported by the EC FP7 grant agreements n. 231855 (sFly) and n.266470 (myCopter)

**Abstract**—In this workshop, we share our experience on autonomous flights with a MAV using a monocular camera as the only exteroceptive sensor. We developed a flexible and modular framework for realtime onboard MAV navigation fusing visual data with IMU at a rate of 1 kHz. We aim to provide a detailed insight of how our modules work and how they are used in our system. Using this framework, we achieved unprecedented MAV navigation autonomy, with flights of more than 350 m length at altitudes between 0 m and 70 m, in previously unknown areas while global signals as GPS were not used.

The great advances in Micro Aerial Vehicle (MAV) technology of about 1.5 kg, have sparked great research interest with the prospect of employing them in search and rescue or inspection scenarios. As a result, we have seen some truly impressive systems showcasing aggressive flight maneuvers such as multiple flips [3] and docking to an oblique wall [4]. It is worth noting however, that all these systems can only operate within the coverage area of sophisticated external tracking systems (e.g. Vicon). It is only very recently that methods without such constraints started emerging, performing real-time, onboard MAV state estimation, pushing towards deployable MAVs in general environments. Given the computational constraints, most MAV navigation systems still struggle with long-term flight autonomy in large and unknown (outdoor) environments where global positioning signals are unreliable or unavailable, limiting the autonomy of MAVs and thus, their usability in real scenarios. Following this demand, we developed a powerful framework achieving unprecedented MAV navigation autonomy, with flights of more than 350 m length at altitudes between 0 m and 70 m, in previously unknown areas while global signals as GPS were not used.

In this workshop, we share our experience on the development of our framework for real-time onboard MAV navigation, with a particular focus on the technical challenges we have been facing. With the most important parts of this framework just made available to the wider Robotics community, we aim to provide a detailed insight of how these modules work and how they are used in our system.

We choose a minimal sensor suite onboard our MAV, consisting of a monocular camera for sensing the scene and an Inertial Measurement Unit (IMU) providing estimates of the MAV acceleration and angular velocity. While this choice implies low power consumption and wide applicability of algorithms (e.g. same sensor suite exists in smart phones),



Fig. 1. The MAV operating in a disaster area, only using measurements from a monocular camera and IMU for localization

the challenges within the estimation processes are great. Drawing inspiration from the latest advances in monocular Simultaneous Localization And Mapping (SLAM) we develop a visual odometry algorithm [6, 8], able to provide ego-motion and scene estimation at constant complexity which is key for autonomous long-term missions.

As the visual pose estimates are prone to drift and can recover position only up to a relative scale, we take advantage of the complementary nature of visual and inertial cues to fuse the unsynchronized sensor readings within an Extended Kalman Filter (EKF) framework [7] for robust MAV state estimation. The 6 DoF MAV pose and velocity are estimated, as well as the visual scale and the intra- and inter-sensor calibration parameters between the camera and the IMU. This scheme enables automatic and online sensor calibration during the mission, rendering our system truly power-on-and-go, while long-term navigation in large areas becomes possible, despite the drift in the visual estimates [5] (which most systems suffer from). Finally, with all computation running onboard, we eliminate the otherwise unrealistic reliance on a permanent communication link to a ground station, bringing MAV navigation closer to usability in real-life scenarios.

We use a prototype of the “FireFly” hexacopter from Ascending Technologies, with similar IMU and soft-/hardware interfaces to the “Pelican”, but improved vibration damping for the sensors and failure tolerance of one rotor. As the most time-critical processes, position control [1, 8] and state prediction run at 1 kHz on a user-programmable microcontroller directly on the flight control hardware of the MAV, enabling immediate reactions on disturbances. The remaining

processes run on an onboard 1.86 GHz Core2Duo embedded computer. Our framework uses the ROS middleware and runs on a standard Ubuntu operating system, facilitating the development of new algorithms. The current implementation, uses only 60 % of one core of the Core2Duo processor at 30 Hz, leaving enough resources for future higher-level tasks. As a reference, the same implementation on an Atom 1.6 GHz single-core computer runs at 20 Hz using 100 % of the processing power.

Our framework has been thoroughly tested under a variety of challenging conditions, exhibiting robustness in the presence of wind gusts, strong light conditions causing saturated images and large scale changes of 0-70 m flight altitude, permitting vision-based landing (Fig. 2). Successful tracking in such scenarios demonstrates the power our robust sensor fusion and estimation approach, as these are often fatal in current SLAM systems. With our setup we were able to perform exploration flights of more than 350 m length (Fig. 3) resulting in an overall position drift of only 1.5 m. Explicitly addressing the computational restrictions typical in MAVs, our distributed processing approach across different units, ensures position control at 1 kHz and enables dynamic trajectory flights with 2 m/s track speed. Finally, the continuous online estimation of calibration and visual drift states, enables long term-navigation and eliminates the need of tedious pre-mission (re-)calibration procedures. On this basis, we believe that our system constitutes a milestone for vision-based MAV navigation in large, unknown and GPS-denied environments providing a reliable basis for further research towards complete missions of search and rescue or inspection scenarios, even with multiple MAVs [2].

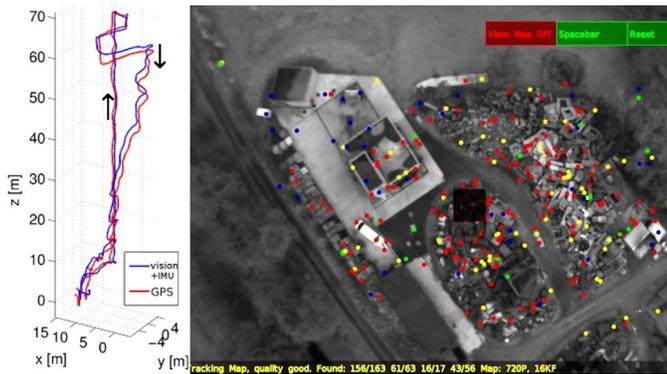


Fig. 2. Altitude test: At  $z \approx 8$  m we switch to vision-based navigation (using sensor feeds from a monocular camera and an IMU only) and fly up to 70 m without intended lateral motion. Then, we let the helicopter descend while applying lateral motion to show the maneuverability during this drastic height change. At the end of the trajectory in the graph, successful vision-based landing is shown, still using only the sensor feeds from the camera and the IMU.

#### WEBLINKS

The software used to achieve the above results has been made publicly available as ROS Packages:

- <http://www.asl.ethz.ch/research/software> (Overview)
- [http://www.ros.org/wiki/ethzasl\\_ptam](http://www.ros.org/wiki/ethzasl_ptam)
- [http://www.ros.org/wiki/ethzasl\\_sensor\\_fusion](http://www.ros.org/wiki/ethzasl_sensor_fusion)



Fig. 3. Trajectory flown in a disaster area. After a short initialization phase at the start, vision-based navigation (blue) was switched on for successful completion of a more than 350 m-long trajectory, until battery limitations necessitated landing. The comparison of the estimated trajectory with the GPS ground truth (red) indicates a very low position and yaw drift of our real-time and onboard visual SLAM framework.

- [http://www.ros.org/wiki/asctec\\_mav\\_framework](http://www.ros.org/wiki/asctec_mav_framework)

#### REFERENCES

- [1] M. W. Achtelik, M. C. Achtelik, S. Weiss, and R. Siegwart. Onboard IMU and Monocular Vision Based Control for MAVs in Unknown In- and Outdoor Environments. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011.
- [2] M. W. Achtelik, S. Weiss, M. Chli, F. Dellaert, and R. Siegwart. Collaborative stereo. In *Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS)*, September 2011.
- [3] Sergei Lupashin, Angela Schöllig, Michael Sherback, and Raffaello D'Andrea. A simple learning strategy for high-speed quadcopter multi-flips. In *ICRA*, pages 1642–1648, 2010.
- [4] D. Mellinger, N. Michael, and V. Kumar. Trajectory generation and control for precise aggressive maneuvers with quadrotors. In *Proceedings of the International Symposium on Experimental Robotics*, Dec 2010.
- [5] S. Weiss and R. Siegwart. Real-time metric state estimation for modular vision-inertial systems. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2011.
- [6] S. Weiss, D. Scaramuzza, and R. Siegwart. Monocular-SLAM based navigation for autonomous micro helicopters in GPS-denied environments. *Journal of Field Robotics*, 28(6):854–874, 2011.
- [7] S. Weiss, M. W. Achtelik, M. Chli, and R. Siegwart. Versatile Distributed Pose Estimation and Sensor Self-Calibration for an Autonomous MAV. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2012.
- [8] S. Weiss, M. W. Achtelik, S. Lynen, M. Chli, and R. Siegwart. Real-time Onboard Visual-Inertial State Estimation and Self-Calibration of MAVs in Unknown Environments. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2012.